Figure 1A **Nucleotide sequence of inserted environmental DNA**

```
TCTCATTTAG TTTGACTGAA ATACCTACTG TGCCACAAAG TAAAGTTAAA          50
CTGACGAATG TGGAACGAAT CACTTAATGG TTCTAGCATA GATAACGAAA         100
GATGAACACG TTCAAAGTTC GCCACTCTTT TGAAAGAGGG TGAACTTTTT         150
TTGTGACAAG AAAGGGTGTT AAATGAAGAT CAAAGCTAAA CAAGATGAGT         200
AACGTTTCTT TTCTTTTTTA TAGAGTGAGT TAGTATATGA GTCCCTTATA         250
AATTTCTAGA CTGTTATTTT AAATAATTGA ATGACTCAGT CACCATTAAG         300
TTTTCAACAC CCATAAGCGA CGTTTGAAGA TCTAATGATG CGAGAGGTTT         350
TATCACTTTG GAGCGGAAGA TCACTGTAGG ACTCGTTTTA TATGGTGAAC         400
TTGGTGTTAT TGTGTATTTA AAAGAAAGGG AAACGAAAAA AATGGTTAAA         450
TTAGAAAGAG GCTATTACAG AGAGGAGAAC AAACAATGAA CGTAACACTT         500
GAAGTGACAT ACTGCACGAC TAAAGGTATT CGAACAACCT TTCATTCAGA         550
AGGTATGGAG GCCGAAAAAG CAATTACCAT CGCAGAAGAT TTTCAGCGGA         600
CAGGACGGAT AAAACAGATC GTCTTTAGAG ATGAGCGTGA TAGTCCGTGG         650
ACGTTAAAAG AACTTAAAAG ATTTTTAGAA GAGATTAAAA CGGAGCCGCA         700
TCATCTCTCT GTGTATTTTG ATGGGGGATT TGATTTGGAG ACACAACGAT         750
CTGGTCTTGG GTGTGATTTA TTATGAACAA AATGACACGT CTTATCGGGT         800
GAGAAGAAAC GCTACCGTGG CGTCATTGAC ATCGAATAAC GAAGCAGAAT         850
ATGCCGCTTT ACATTTAGGA CTTAAAGAAC TTGAAGGGAT CGGTGCGCAT         900
CATCTACCTA TCACTATTTA CGGTGATTCT CAAGTTGTGA TCAATCAGTT         950
AAAAGGAGAA TGGGCGTGTA TGGAGGAGGT GTTAAATAAA TGGGCTGACC        1000
GTATTGATCA GCATTTAGCT AAATTAGGCA TGACCGCTAC TTATAAGTTA        1050
ATCCCCCGTA AAGAAAACCG TGAAGCAGAT CAACTGGCTA CACAAGCGTT        1100
AAACGGGCAA GAAATTATAA GTCAACGTGA TGTCAGTGAG CGTGGTGCAG        1150
ATTAGTCTGC ACCCGCATAA AAGTTAACGT ATATAGAAGT GGATGGGGAT        1200
TAAAGGAACG TCATTCACTC TAAGCAAGCG TTGCGACAGC AAAAAAGAAA        1250
CATATAAGGT TTTTCTGAGC TACTATCTAT ACAAATAGCC AAGTGGCAGT        1300
TAAGCTCTTA CCTCATCAAG TTTTTGACTA CCAGTCTTCC ACTCCTACTT        1350
TCACCTATAT AAATTGGTTC CTTTTTTGTT AATAATCACT AATTTTGACG        1400
GTATTTTTTA ATAGAAATAT ATGCTAGATT ATAAACTAGT AACGATGTAG        1450
AAGGTGGTGA TTGACCATAT AAGAAGACTC TTTCAAACCT GGTAGTATCG        1500
CATTAAAAAA TTTGAAAGGT GGAGAGGACA CATGGGTTAT ACCAAAGCGA        1550
AGTGTACGTT GAAAAAAACT GTCTTGTTTG GTTTAATTCT CTGTTTAAGT        1600
GTGTCAATGT TTGTTCCAAT GACATCAGCT GAAGATGTCA CTTCGTCACA        1650
GTTGGATATT CACTCCTATG TAGCTGACAT GCAGCCTGGC TGGAATTTAG        1700
GAAATACGTT TGACGCTGTT GGAGATGATG AAACAGCGTG GGGGAATCCT        1750
CGTGTAACAA GAGAGTTAAT AAAAACGATT GCTGATGAAG GGTATAAAAG        1800
CATTCGTATC CCAGTGACAT GGCAAAATCA AATGGGTGGT TCTCCAGATT        1850
ATACGATAAA TGAAGATTAT ATCAATCGGG TGGAGCAAGC GATAGATTGG        1900
GCGTTGGAGG AAGACTTATA TGTGATGTTA AATGTGCATC ATGACTCATG        1950
GCTGTGGATG TATGATATGG AACATAACTA TGATGAGGTC ATGGCAAGAT        2000
ATACAGCTAT TTGGGAACAA TTGTCGGAAA AATTCAAAAG CCACTCCCAT        2050
AAGTTGATGT TTGAGAGTGT CAATGAGCCT AGGTTTACGC AGGAGTGGGG        2100
AGAGATTCAA GAAAATCATC ATGCTTACTT AGAAGATTTA AATAAGACGT        2150
TCTATTATAT TGTCAGAGAG TCAGGAGGCA ATAATGTGGA GCGCCCTTTA        2200
GTATTGCCTA CGATAGAAAC AGCCACGTCT CAGGATTTAC TAGATCGCTT        2250
GTATCAAACA ATGGAAGACT TGGATGATCC TTATTTAATT GCCACGGTGC        2300
ATTATTATGG CTTCTGGCCA TTTAGTGTCA ATATAGCAGG GTACACTCAT        2350
TTTGAACAGG AAACACAACA AGATATTATA GACACCTTTG ACCGTGTTCA        2400
TAACACATTT ACAGCGCGTG GTGTCCCAGT TGTATTAGGC GAATTCGGTT        2450
```

Figure 1B

```
TGTTAGGCTT TGACAAAAGT ACGGATGTGA TTCAGCAAGG GGAGAAATTA          2500
AAGTTTTTTG AGTTTCTCAT CCATCATCTC AATGAACGTG ATATAACCCA          2550
TATGTTATGG GATAACGGCC AGCATTTAAA TCGAGAAACT TATGCATGGT          2600
ATGATCAAGA ATTTCATGAC ATATTAAAAG CGAGTTGGGA GGGGCGTTCT          2650
GCTACAGCAG AGTCTAATTT GATTCATGTG AAGGACGGAA AGCCAATTAG          2700
AGATCAAGAT ATACAGCTTT ACTTAAACGG AAATGAGCTA ACAGCCTTAC          2750
AGGCAGGGGA GGAATCGCTT GTTCTAGGAG AGGATTATGA ACTAGCAGGA          2800
GGCGTATTAA CGCTAAAAGC GGACACCCTC ACAAGACTAA TTACCCCTGG          2850
TCAATTAGGA ACCAATGCAG TCATCACAGC ACAATTTAAT TCTGGAGCAG          2900
ACTGGCGTTT TCAATTACAG AATGTGGACG TGCCAACGGT CGAAAATACA          2950
GATGGCTCAA CATGGCATTT TGCGATCCCT ACCCATTTTA ATGGTGATAG          3000
TCTTGCGACG ATGGAAGCTG TTTATGCAAA CGGAGAATAT GCTGGGCCGC          3050
AAGATTGGAC GTCATTTAAA GAATTTGGCG AGGCGTTTTC TCCTAATTAC          3100
GCCACAGGGG AAATTATTAT ATCAGAAGCC TTCTTTAACG CGGTACGGGA          3150
TGATGATATC CATTTAACAT TTCATTTTTG GAGCGGAGAG ACGGTGGAAT          3200
ATACCTTACG TAAAAATGGC AATTATGTTC AAGGTAGACG GTAACATGAT          3250
TTTAATTAAT AGATAAACCA GCCTACCTAT CGTTTTTGGA AGAAGGCAAA          3300
CGAATCTCAT CTTACCAACA CCGTGCTTTA GAACTTTAGA AGTGACGGTG          3350
TTTTTTAAGA CATGAGGAGA AGCAATCCTC TATCAACAGT CACCAATTTT          3400
TATTCAGGAG GTGTCAAGTT ATCTAACGTT CTATGAATGC ATATAGTCTC          3450
TGACGAATAA ACATAGTTAA AAAGAAGTGA GCCTAGTCCC CGAGGGGAAG          3500
GGGATAATGC CAACGTATTG GATTAAAGTA CCTTCTTGAT AAAAAGAAAG          3550
GGTTTTCAAG AGATGGAAAT GGGCTCGTTT GTTATACTTT AATTACGCCT          3600
TGGAACGTCA TTTTGGCGGT GCTATTTAGT AAGATGGCTG ACATCATAAA          3650
AGAGGAGTGG GTTCGATGGC TTTAATTCAA TTAAGCTTTA AATCACGAGC          3700
ATTAATGTTG CAAACCTCTG TCAATGTTTT ATTACCGGTG GGAATGAATG          3750
CTGTAGATTT TACACCAAGT GATGATTTTT CTTATGTTAC TGACCCTTTT          3800
CCTGTCCTAT ATCTTTTGCA TGGTGCAACT GATGATTATT CAGCATGGCT          3850
ACGTCTGTCC TCTATTGAAC GATATGCTGA AGAAAAAAAA TTGGCGGTCG          3900
TCATGCCAAA TGCTGATATG AGTGCGTATA CGGATATGGT ACATGGCCAT          3950
CGTTACTGGA CGTATATTAG TAAGGTGCTG CCTGAGTTTA TGAGAGCAAC          4000
TTTTCCTATT TCTCAGCACC GTGAAGACAC CTTTGCAGCT GGTCTGTCTA          4050
TGGGAGGATA CGGGGCTTTT AAATTGGCGC TGCGGCAACC GGAACGCTTC          4100
GCTGCAGCTG TGTCATTATC TGGTGCAGTT GATATGAGAG AAGCAAGTCA          4150
ACCAGACTCC CTATTTGTAA ATGCCTTTGG TGAAGGGACG AAAATCGCAG          4200
GGACA                                                          4205
```

Figure 2 **ORF Nucleotide sequence of cellulase gene**

```
ATGGGTTATA CCAAAGCGAA GTGTACGTTG AAAAAAACTG TCTTGTTTGG          50
TTTAATTCTC TGTTTAAGTG TGTCAATGTT TGTTCCAATG ACATCAGCTG         100
AAGATGTCAC TTCGTCACAG TTGGATATTC ACTCCTATGT AGCTGACATG         150
CAGCCTGGCT GGAATTTAGG AAATACGTTT GACGCTGTTG GAGATGATGA         200
AACAGCGTGG GGGAATCCTC GTGTAACAAG AGAGTTAATA AAAACGATTG         250
CTGATGAAGG GTATAAAAGC ATTCGTATCC CAGTGACATG GCAAAATCAA         300
ATGGGTGGTT CTCCAGATTA TACGATAAAT GAAGATTATA TCAATCGGGT         350
GGAGCAAGCG ATAGATTGGG CGTTGGAGGA AGACTTATAT GTGATGTTAA         400
ATGTGCATCA TGACTCATGG CTGTGGATGT ATGATATGGA ACATAACTAT         450
GATGAGGTCA TGGCAAGATA TACAGCTATT TGGGAACAAT TGTCGGAAAA         500
ATTCAAAAGC CACTCCCATA AGTTGATGTT TGAGAGTGTC AATGAGCCTA         550
GGTTTACGCA GGAGTGGGGA GAGATTCAAG AAAATCATCA TGCTTACTTA         600
GAAGATTTAA ATAAGACGTT CTATTATATT GTCAGAGAGT CAGGAGGCAA         650
TAATGTGGAG CGCCCTTTAG TATTGCCTAC GATAGAAACA GCCACGTCTC         700
AGGATTTACT AGATCGCTTG TATCAAACAA TGGAAGACTT GGATGATCCT         750
TATTTAATTG CCACGGTGCA TTATTATGGC TTCTGGCCAT TTAGTGTCAA         800
TATAGCAGGG TACACTCATT TTGAACAGGA AACACAACAA GATATTATAG         850
ACACCTTTGA CCGTGTTCAT AACACATTTA CAGCGCGTGG TGTCCCAGTT         900
GTATTAGGCG AATTCGGTTT GTTAGGCTTT GACAAAAGTA CGGATGTGAT         950
TCAGCAAGGG GAGAAATTAA AGTTTTTTGA GTTTCTCATC CATCATCTCA        1000
ATGAACGTGA TATAACCCAT ATGTTATGGG ATAACGGCCA GCATTTAAAT        1050
CGAGAAACTT ATGCATGGTA TGATCAAGAA TTTCATGACA TATTAAAAGC        1100
GAGTTGGGAG GGGCGTTCTG CTACAGCAGA GTCTAATTTG ATTCATGTGA        1150
AGGACGGAAA GCCAATTAGA GATCAAGATA TACAGCTTTA CTTAAACGGA        1200
AATGAGCTAA CAGCCTTACA GGCAGGGGAG GAATCGCTTG TTCTAGGAGA        1250
GGATTATGAA CTAGCAGGAG GCGTATTAAC GCTAAAAGCG GACACCCTCA        1300
CAAGACTAAT TACCCCTGGT CAATTAGGAA CCAATGCAGT CATCACAGCA        1350
CAATTTAATT CTGGAGCAGA CTGGCGTTTT CAATTACAGA ATGTGGACGT        1400
GCCAACGGTC GAAAATACAG ATGGCTCAAC ATGGCATTTT GCGATCCCTA        1450
CCCATTTTAA TGGTGATAGT CTTGCGACGA TGGAAGCTGT TTATGCAAAC        1500
GGAGAATATG CTGGGCCGCA AGATTGGACG TCATTTAAAG AATTTGGCGA        1550
GGCGTTTTCT CCTAATTACG CCACAGGGGA AATTATTATA TCAGAAGCCT        1600
TCTTTAACGC GGTACGGGAT GATGATATCC ATTTAACATT TCATTTTTGG        1650
AGCGGAGAGA CGGTGGAATA TACCTTACGT AAAAATGGCA ATTATGTTCA        1700
AGGTAGACGG TAA                                               1713
```

Figure 3 **Amino acid sequence of BagCel cellulase**

```
MGYTKAKCTL KKTVLFGLIL CLSVSMFVPM TSAEDVTSSQ LDIHSYVADM          50
QPGWNLGNTF DAVGDDETAW GNPRVTRELI KTIADEGYKS IRIPVTWQNQ         100
MGGSPDYTIN EDYINRVEQA IDWALEEDLY VMLNVHHDSW LWMYDMEHNY         150
DEVMARYTAI WEQLSEKFKS HSHKLMFESV NEPRFTQEWG EIQENHHAYL         200
EDLNKTFYYI VRESGGNNVE RPLVLPTIET ATSQDLLDRL YQTMEDLDDP         250
YLIATVHYYG FWPFSVNIAG YTHFEQETQQ DIIDTFDRVH NTFTARGVPV         300
VLGEFGLLGF DKSTDVIQQG EKLKFFEFLI HHLNERDITH MLWDNGQHLN         350
RETYAWYDQE FHDILKASWE GRSATAESNL IHVKDGKPIR DQDIQLYLNG         400
NELTALQAGE ESLVLGEDYE LAGGVLTLKA DTLTRLITPG QLGTNAVITA         450
QFNSGADWRF QLQNVDVPTV ENTDGSTWHF AIPTHFNGDS LATMEAVYAN         500
GEYAGPQDWT SFKEFGEAFS PNYATGEIII SEAFFNAVRD DDIHLTFHFW         550
SGETVEYTLR KNGNYVQGRR                                          570
```